Reserve Bank of Australia

RESEARCH
DISCUSSION
PAPER

Forecasting with Factors: The Accuracy of Timeliness

Christian Gillitzer and Jonathan Kearns

RDP 2007-03

FORECASTING WITH FACTORS: THE ACCURACY OF TIMELINESS

Christian Gillitzer and Jonathan Kearns

Research Discussion Paper
2007-03

April 2007

Economic Analysis Department
Reserve Bank of Australia

Thanks to Christopher Kent, Kristoffer Nimark, Anthony Richards and Louise Wilkinson for helpful comments and suggestions. The views expressed in this paper are those of the authors and do not necessarily reflect those of the Reserve Bank of Australia.

Authors: gillitzerc or kearnsj at domain rba.gov.au

Economic Publications: ecpubs@rba.gov.au

Abstract

This paper demonstrates that factor-based forecasts for key Australian macroeconomic series can outperform standard time-series benchmarks. In practice, however, the advantages of using large panels of data to construct the factors typically comes at the cost of using less timely series, thereby delaying when the forecasts can be made. To produce more timely forecasts it is possible to use a narrower data panel, though this will possibly result in less accurate factor estimates and so less accurate forecasts. We demonstrate this trade-off between accuracy and timeliness with out-of-sample forecasts. With the exception of only consumer price inflation, the forecasts do not become less accurate as they utilise less information by excluding less timely series. So while factor forecasts have large data requirements, we show that these should not prevent their practical use when timely forecasts are needed.

JEL Classification Numbers: C53, E27, E37
Keywords: forecasting, factor models, Australia

Table of Contents

FORECASTING WITH FACTORS: THE ACCURACY OF TIMELINESS

Christian Gillitzer and Jonathan Kearns

1. Introduction

Dynamic factor models, which enable a large number of economic time series to be combined, have been shown to frequently produce more accurate forecasts than standard time-series benchmarks, structural models and even official forecasts, especially at horizons of between one and two years.[1] There are good reasons for this. Using a large number of series can produce a measure of current economic activity that contains less noise than the individual series used in more traditional forecasting frameworks. Further, using many series will typically make the forecasts more robust to structural change in individual explanatory variables and possibly to structural change among the relationships between economic variables. However, there are potential costs to using many series when forecasting. One which has received little attention in the literature is that a larger panel of series will contain less timely series and so can only be used to produce forecasts with a considerable delay. In this paper we examine how forecast accuracy changes when factor forecasts are made using a more timely panel which necessarily contains less information. This is an important trade-off for policy-makers: more accurate forecasts will help policy-makers, but so too will more timely forecasts.

In this paper we examine this trade-off between accuracy and timeliness for eight key Australian macroeconomic series. More timely factor forecasts can be made by using a narrower panel that effectively excludes the information in series with late release dates. Using a narrower panel is likely to reduce the precision of the

1 This finding has most frequently been demonstrated for inflation: see Angelini, Henry and Mestre (2001) for the euro area; Stock and Watson (1999), Brave and Fisher (2004) and Gavin and Kliesen (2006) for the United States; Gosselin and Tkacz (2001) for Canada; and Moser, Rumler and Scharler (2004) for Austria. Other studies suggest that these results generalise to real variables and interest rates: see Banerjee, Marcellino and Masten (2005) for five new European Union member states; Matheson (2005) for New Zealand; Stock and Watson (2002b) for the United States; and Artis, Banerjee and Marcellino (2005) for the United Kingdom.

estimates of the factors, but it remains an open question whether the resulting deterioration in forecast accuracy is large. Several papers have found that the deterioration in forecast accuracy when using smaller panels is not large. In this regard, it is relevant that Boivin and Ng (2006) and Watson (2003) find that forecast accuracy does not improve beyond the use of 50–100 series. Similarly, Schneider and Spitzer (2004) found that they needed to restrict the size of the panel to outperform simple benchmarks when forecasting Austrian GDP.

There are, however, other limitations to the use of factor forecasts for policy purposes that we do not address in this paper. One of the most significant of these is that they are reduced-form forecasts based only on contemporaneous information and so they cannot be conditioned on particular assumptions. Notably, the forecasts cannot be conditioned on a specific path of interest rates. This may limit their usefulness in a policy environment, such as a central bank.

The remainder of this paper proceeds as follows. In Section 2 we briefly describe the factor forecasting procedure. We discuss the timeliness of data and the composition of the panel in Section 3. In Section 4 we document the performance of factor forecasts and the empirical trade-off between accuracy and timeliness, before concluding in Section 5.

2. Factor Models and Timely Forecasting

The process of using factors to forecast can be broken down into two steps. First, a panel of data is used to estimate the factors. Second, these factors are used to produce out-of-sample forecasts for the series in question. In this section we explain the estimation of the factors and the forecasting equation, though because these techniques have been discussed elsewhere, for example Boivin and Ng (2005), we only provide a brief exposition.

2.1 Estimating the Factors

The data used to estimate the factors are assumed to have an approximate factor representation, given by Equation (1),

$$x_{it} = \lambda_{i0} f_t + \lambda_{i1} f_{t-1} + \ldots + \lambda_{is} f_{t-s} + \varepsilon_{it} \tag{1}$$

where x_{it} is the time t observation of series i of the data panel, λ_{ij} is a vector of series-specific factor loadings for lag j of the factors, f_t is a vector of q factors common across all series and ε_{it} is a series-specific idiosyncratic error term (which may be weakly correlated across time and series).

To estimate the factors we use the method demonstrated by Stock and Watson (1999; 2002a; 2002b). Given the simple representation presented by Equation (1), the factors and loadings can be estimated by calculating the principal components of the data panel. If Equation (1) includes lags, then the factors are estimated by principal components of a matrix that augments the data panel with lags of the data panel; for example, if the matrix of the x_{it} is denoted by X and one lag is included, then the principal components are calculated from the matrix that concatenates the matrix of data from time 0 to $T-1$ with that from time 1 to T, that is, $X_{[0,T-1]}|X_{[1,T]}$.

2.2 Forecasting

The second step involves including the estimated factors in a forecasting regression. Various specifications of the forecasting equation have been used in the literature. We focus on a simple one that can easily be used for our iterative process in which we use panels with varying numbers of series, and has typically been found to produce forecasts that are at least as good as other specifications. The series being forecast, y_{t+h}^h, is regressed on current and lagged estimates of the factors, $\widehat{f}_t$, and lags of itself, as in Equation (2).

$$y_{t+h}^h = \alpha_h + \beta_h(L)\widehat{f}_t + \gamma_h(L)y_t + \varepsilon_{t+h}^h \tag{2}$$

We use the notation y_{t+h}^h to show that the dependent variable is the h-period percentage change of the series being forecast (the one exception is the unemployment rate for which we use the h-period ahead level). The lags on the right-hand side of the equation, $\gamma_h(L)y_t$, are one-period percentage changes (or the level for the unemployment rate). All percentage changes are approximated by log differences. As Boivin and Ng (2005) demonstrate, including autoregressive terms (lags of the dependent variable) along with the factors is equivalent to forecasting using the factors and the lagged idiosyncratic terms, ε. After estimating Equation (2), the forecast is then generated using the parameter estimates.

In Appendix B we compare the performance of the forecasts generated using this methodology with forecasts that use the 'dynamic' factors estimated by the technique of Forni *et al* (2005) and a non-parametric technique developed by these authors. The simpler methodology employed here is shown to produce forecasts that are at least as good as those using these more complex procedures. This result is in line with other studies, for example Boivin and Ng (2005).

2.3 Timeliness of the Data Panel

The time it takes for data for a given quarter to be released varies across different macroeconomic series. Consequently, as more time passes since the last quarter of the in-sample period, which is the base quarter from which the forecasts are made, an increasing number of series will have an observation for that quarter. To examine the trade-off between forecast accuracy and timeliness we estimate the factors and forecasts recursively using data panels expanded to include the increasing number of series that become available as the number of days since the end of the base quarter increases. This enables us to examine how forecast accuracy changes as we wait for more series to become available so that a broader panel can be used to estimate the factors.

In order to incorporate the lagged information that is contained in these series for which the base quarter's data are not available, we use a pseudo stacked panel that includes the one-period lag of all series and the contemporaneous values for the series that have been released.[2] For example, if the panel X contains 50 series, but only 20 of these have been released one month after the end of the base quarter, then the full matrix from which we estimate the factors at this time contains 70 series and is $X_{[0,T-1]}|\widetilde{X}_{[1,T]}$, where $\widetilde{X}_{[1,T]}$ is the matrix representing the panel of these 20 available series. The iterative procedure of expanding the panel breadth starts with the release of the first series, when $\widetilde{X}_{[1,T]}$ contains just 1 series, and proceeds until all the series have been released and so $\widetilde{X}_{[1,T]}$ contains all 50 series, and is the same as $X_{[1,T]}$.

2 An alternative approach would be to impute the missing observations using the expectations maximisation algorithm described by Stock and Watson (2002b). However, this technique does not tend to produce reliable estimates for missing observations.

3. The Composition of Forecast Models

In this study we use a quarterly panel because most of the key Australian macroeconomic time series that one would wish to forecast are available at a quarterly frequency, including the consumer price index (CPI). Further, over the long sample period used in this study, there are relatively few monthly series available for Australia that comprise a representative cross-section of the economy.

Table 1 summarises the composition of our data panel according to the type of economic series. The panel includes 53 series in total and spans the period 1960−2005 (see Appendix A for a full list of the data series). The composition of the data panel is crucial for ensuring that the estimated factors are representative of the aggregate economy. As Gillitzer, Kearns and Richards (2005) discuss, a panel that has a balance of series representing different aspects of the economy is more likely to produce factors that reflect the entire economy. For instance, if the data panel contained a disproportionate number of labour market series, the estimated factors would more closely approximate the state of the labour market than the aggregate economy. The data panel is also constrained by our desire to have a long data sample to ensure that the results are robust to structural change and the state of the economic cycle. Because real-time data are not available for all series in our panel we use final vintage data.

Table 1: Composition of the Data Panel

Type of series	Number of series
National accounts	21
Employment	8
Industrial production	4
Building and capital expenditure	5
Internal trade	1
Overseas trade and current account	5
Prices	5
Financial data	4
Total	53

While the data panel contains fewer series than used in similar studies, this need not result in less accurate forecasts. Broader data panels typically include many highly disaggregated series, such as regional series or subcomponents of the series that we have used. These may contain more noise and correlated idiosyncratic dynamics, thereby potentially reducing the information content of the estimated factors (see Boivin and Ng 2006).

The factor models we estimate require the data to be in stationary form, which for most series, such as GDP, we achieve by taking a log-difference, leaving the data in approximate percentage change form. Series such as business surveys and the unemployment rate are already stationary and so require no transformation.

We forecast eight important macroeconomic series representing different aspects of the economy: growth in GDP, non-farm GDP, private final demand, household final consumption expenditure, employment, the number of building approvals, CPI inflation and the unemployment rate.

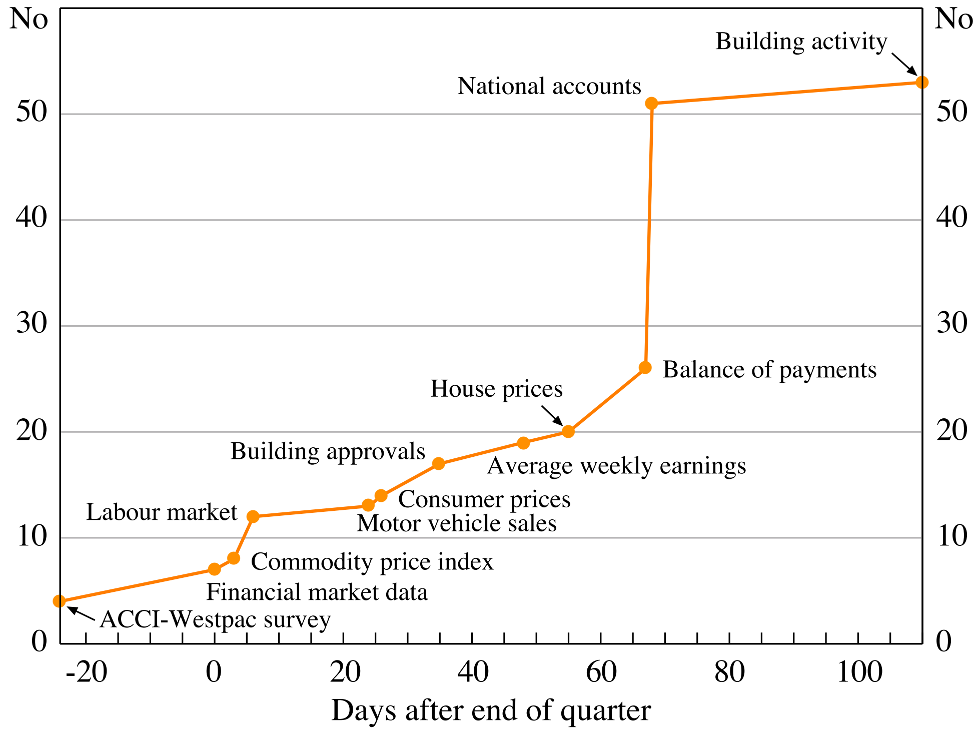
The publication lag for the data in the panel is shown in Figure 1. Using the release dates for 2006:Q1, it shows the number of series in the panel that have been published at a given time relative to the end of the quarter that those data cover. While the release dates for series in the panel may have changed over the sample, the timing depicted in Figure 1 presents the data availability constraints currently faced by forecasters, and is representative of the order in which series have been released over the long sample period used in this study.

The most timely series in the panel are from the ACCI-Westpac survey of manufacturers, and arrive around 20 days before the end of the relevant quarter. The financial market data in the panel are available within a few days of the end of the quarter, followed about a week later by the labour market data, while many real and nominal series become available between approximately 3–8 weeks after the end of the quarter.[3] Some of the least timely series in the panel are the balance of payments and national accounts data, which are available about 10 weeks after the end of the quarter.

3 Financial market data covering part of the quarter would be available contemporaneously, but to keep the exercise tractable, we only consider the data covering the full quarter.

Figure 1: Number of Data Series in Panel

Days after the end of the quarter



Many similar types of economic series – for example, financial market series or labour market series – are typically released at about the same time. Consequently, prior to the point at which all the series become available, the subset of the data panel with observations for the base quarter will not be representative of the full panel as it will exclude those groups of related series yet to be released. Given the constraints of release dates, this is unavoidable. If national accounts series are not released until 10 weeks after the end of the base quarter, then the data panel used 9 weeks after the end of the quarter cannot include the base quarter's data for any of these series. However, our choice of series ensures that each subset of series that include the base quarter's data is as representative as possible, given data availability. For example, while the group of series with data for the base quarter available 10 days after the end of the quarter will not contain production data, it will have a reasonable balance of survey, financial market and labour market series.

4. Forecast Accuracy

We produce out-of-sample forecasts to assess the accuracy of factor forecasts for our eight macroeconomic series. The factors and forecasting equations are estimated over an in-sample period (initially 1960:Q3–1970:Q1). The estimated factors and forecasting equations are then used to produce forecasts for horizons of two, four and eight quarters from the end date of this in-sample period. The in-sample period is then lengthened by one quarter at a time with forecasts for the three horizons produced at each step. The last forecast is for 2005:Q4, giving a sequence of almost 140 forecasts for each horizon.

Rewriting Equation (2) shows the most general specification of the forecasting equation estimated in this paper:

$$y^h_{t+h} = \widehat{\alpha}_h + \sum_{j=1}^{m} \widehat{\beta}_{hj} \widehat{f}_{t-j+1} + \sum_{j=1}^{p} \widehat{\gamma}_{hj} y_{t-j+1} + \varepsilon^h_{t+h} \tag{3}$$

where $\widehat{f}_t$ is the vector of q estimated factors at time t, and y_t is the forecast series, which is in log-difference form for all forecast series except the unemployment rate which is in level form.

Because there are numerous ways for determining the number of factors (q) and the number of lags of the factors (m) and the forecast variable (p) included in Equation (3), there are many possible variants of the forecasting equation. Existing studies, such as Stock and Watson (2002a) and Gavin and Kliesen (2006), estimate and present many versions of the forecasting equation. While we also examined the forecasting performance of numerous versions of Equation (3), so as not to overwhelm the reader we present results for just three variants (although those are representative of the broader results). We present the same specifications for all forecast variables to limit any sense of 'mining' the numerous models to present more favourable results for each series. The full set of results are available from the authors on request.

The first two specifications are simple, containing a fixed number of factors at each forecast iteration. The model denoted F2 includes just two factors, with no lags of the factors or the forecast variable. The model denoted FAR2 includes two factors (but with no lagged factors) and also allows for the inclusion of up to three lags of the forecast variable ($0 \leq p \leq 3$), with the number of lags selected at each iteration using the Bayesian Information Criterion (BIC). We include two factors as they account for about one-quarter of the total variation in the full data panel. The third specification, denoted FAR-BIC, allows the BIC to select both the number of factors (up to six, $1 \leq k \leq 6$) and the number of lags of the forecast variable (up to three, $0 \leq p \leq 3$) at each iteration (again, no lags of the factors are included). This model imposes little structure on the forecasting equation, and so is illustrative of the accuracy of out-of-sample forecasts when there is uncertainty about the appropriate model specification.

We evaluate the accuracy of each forecast model by calculating the mean-squared errors (MSE) of the forecasts for each horizon. We also calculate the MSE of forecasts generated by a simple autoregressive model. This is a commonly used benchmark in out-of-sample forecasting exercises, which has been shown to be difficult to beat, nesting both random walk and constant growth forecasts within the specification. We present our forecast accuracy results as the ratio of the MSE from the factor forecasts to the MSE of the autoregressive forecasts. Numbers less than unity indicate that the factor forecast outperforms the benchmark autoregressive forecast.

4.1 Forecast Accuracy for the Full Panel

Before addressing the issue of timeliness, we demonstrate the performance of factor forecasts using the full data panel. Because all series are included, we do not stack the panel with lags of the series and so the panel used to estimate the factors contains 53 series. Table 2 reports the MSE ratios for the three forecasting equation specifications, at forecast horizons of two, four and eight quarters. Robust standard errors, which account for heteroskedasticity and serial correlation of the forecast errors, are reported in parentheses for each MSE ratio.

For the majority of series and forecast horizons, the MSE ratio is less than unity, indicating that the factor-based forecasts outperform benchmark autoregressive forecasts. For example, the MSE ratio of 0.85 for the F2 model of GDP growth at a four-quarter horizon indicates that the factor forecast has a 15 per cent lower MSE than the autoregressive forecast. Although most MSE ratios are not significantly less than unity at the 5 per cent level of significance, just under half are at the 10 per cent level of significance.

Table 2: Forecasting Performance

Ratio of mean-squared forecast error of candidate model to an autoregressive forecast

Model	Horizon					
	Two quarters		Four quarters		Eight quarters	
GDP growth						
F2	0.92	(0.12)	0.85	(0.14)	0.83*	(0.13)
FAR2	0.96	(0.17)	0.89	(0.17)	0.84	(0.14)
FAR-BIC	1.10	(0.14)	1.02	(0.16)	0.78*	(0.14)
Non-farm GDP growth						
F2	0.78**	(0.12)	0.76*	(0.15)	0.73*	(0.17)
FAR2	0.86	(0.15)	0.87	(0.14)	0.74*	(0.17)
FAR-BIC	0.95	(0.14)	0.93	(0.14)	0.78*	(0.15)
Private final demand growth						
F2	0.82*	(0.12)	0.71**	(0.15)	0.77*	(0.15)
FAR2	0.78**	(0.13)	0.67**	(0.17)	0.73*	(0.17)
FAR-BIC	0.81*	(0.13)	0.74*	(0.18)	0.80	(0.16)
Household final consumption expenditure growth						
F2	0.90*	(0.07)	0.69***	(0.11)	0.64**	(0.16)
FAR2	0.97	(0.08)	0.69***	(0.11)	0.65**	(0.16)
FAR-BIC	1.20	(0.26)	0.97	(0.21)	1.00	(0.26)
Employment growth						
F2	0.76**	(0.14)	0.71**	(0.16)	0.77*	(0.17)
FAR2	0.78**	(0.13)	0.76*	(0.17)	0.80	(0.18)
FAR-BIC	0.82*	(0.12)	0.84	(0.14)	0.81	(0.17)
Unemployment rate						
F2	16.02	(4.06)	4.08	(2.65)	1.80	(0.61)
FAR2	0.67**	(0.16)	0.71**	(0.16)	0.76**	(0.13)
FAR-BIC	0.58**	(0.19)	0.70**	(0.17)	0.71**	(0.15)
CPI inflation						
F2	1.78	(0.56)	1.63	(0.40)	1.14	(0.22)
FAR2	0.89	(0.12)	0.84	(0.16)	0.71*	(0.20)
FAR-BIC	1.01	(0.20)	0.78	(0.21)	0.63*	(0.28)
Building approvals growth						
F2	1.09	(0.10)	1.04	(0.09)	0.94	(0.11)
FAR2	1.01	(0.08)	1.04	(0.09)	0.98	(0.12)
FAR-BIC	0.98	(0.09)	0.98	(0.09)	1.01	(0.11)

Notes: Model F2 includes two factors and no lags of the forecast variable. Model FAR2 includes two factors and up to three lags of the forecast variable, selected at each iteration using the BIC ($0 \leq p \leq 3$). Model FAR-BIC uses the BIC to select both the number of factors (up to six, $1 \leq k \leq 6$) and the number of lags of the forecast variable (up to three, $0 \leq p \leq 3$) at each iteration. Numbers in parentheses are robust standard errors calculated using the delta method. Ratios significantly less than 1 at the 1, 5 and 10 per cent confidence levels are indicated by ***, ** and *.

The results in Table 2 show that the MSE ratio is generally lowest for horizons of four and eight quarters (with the unemployment rate a notable exception). This finding that the gains in forecast accuracy are greatest at these longer horizons is consistent with the literature. Note that this does not mean that the factor forecasts are more accurate at longer horizons than at shorter horizons; the absolute MSE of factor forecasts does increase with the horizon of the forecast (not shown). Rather, this result highlights the tendency for factor forecasts to be more accurate relative to the autoregressive forecasts at longer horizons. The improvement in forecasting performance over the autoregressive benchmark is most important from a policy perspective at longer horizons, in part because alternatives – including the use of partial indicators and the importance of recent shocks – are often available for making reasonable short-horizon forecasts.

For the four national accounts series (the first four series in Table 2) and employment, the simple models (F2 and FAR2) which keep the number of factors fixed produce more accurate forecasts than the more complex model which selects the number of factors at each forecast iteration. In general, the simplest model (F2), which maintains the same forecasting equation specification at each iteration, is slightly more accurate. This result contrasts with the poor forecasting performance of the F2 model for consumer prices and the unemployment rate; at all horizons it is less accurate than the benchmark autoregressive forecast. For these two series the inclusion of lags of the forecast variable in the forecasting equation is important for forecast performance as demonstrated by the lower MSE ratios of the FAR2 and FAR-BIC models. The unemployment rate and inflation have both had long cycles and have likely experienced substantial structural change over the 45-year sample. The factors capture the general state of the economy rather than structural change, and so it is necessary to include lags of the forecast variable to account for any structural change.[4] For the building approvals series, all three models produce forecasts that are no better or worse than the benchmark autoregressive forecast. Surprisingly, the current state of the

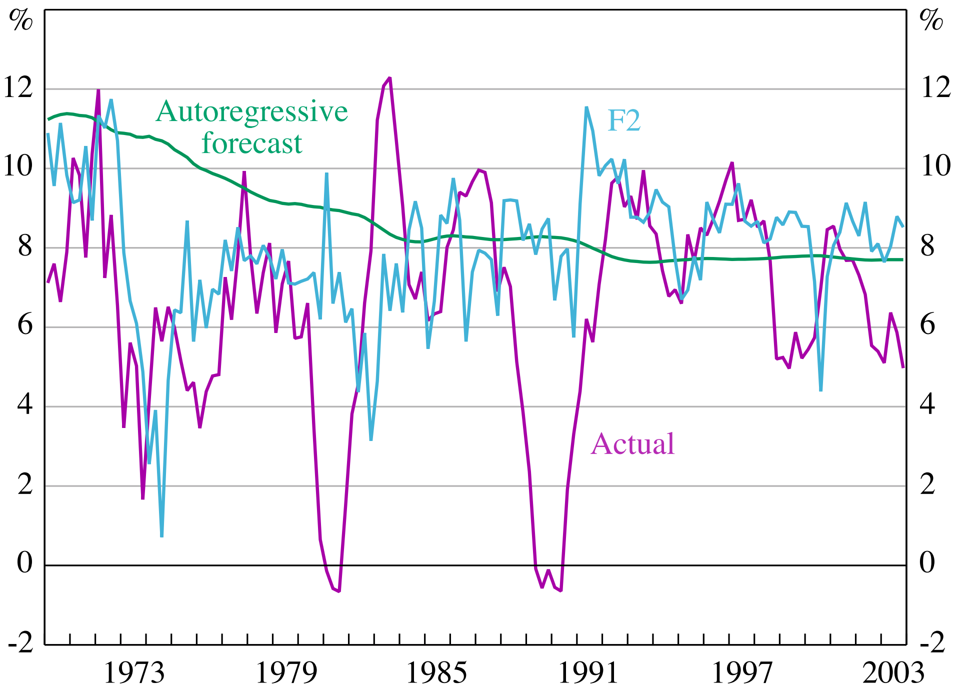
4 Because of these long cycles and likely structural change, the sum of the autoregressive coefficients is close to unity.

overall economy (as indicated by the factors) seems to have little information for forecasting building approvals.[5]

To give an indication of the out-of-sample forecasting performance of the factor-based models, Figures 2 and 3 illustrate respectively the forecasts for eight-quarter growth in non-farm GDP (using the F2 forecast equation specification) and the level of the unemployment rate eight quarters ahead (using the FAR2 specification). The autoregressive forecasts for each series are also shown. For non-farm GDP, the autoregressive forecast chosen by the BIC is a simple average of growth of non-farm GDP in the in-sample period. Clearly, the factor-based forecasts are able to capture a substantial amount of variation in the forecast series over and above the autoregressive forecasts. In contrast, for the unemployment rate the factor forecast is very similar to the autoregressive forecast, demonstrating the importance of lags of the forecast variable in forecasting this series.

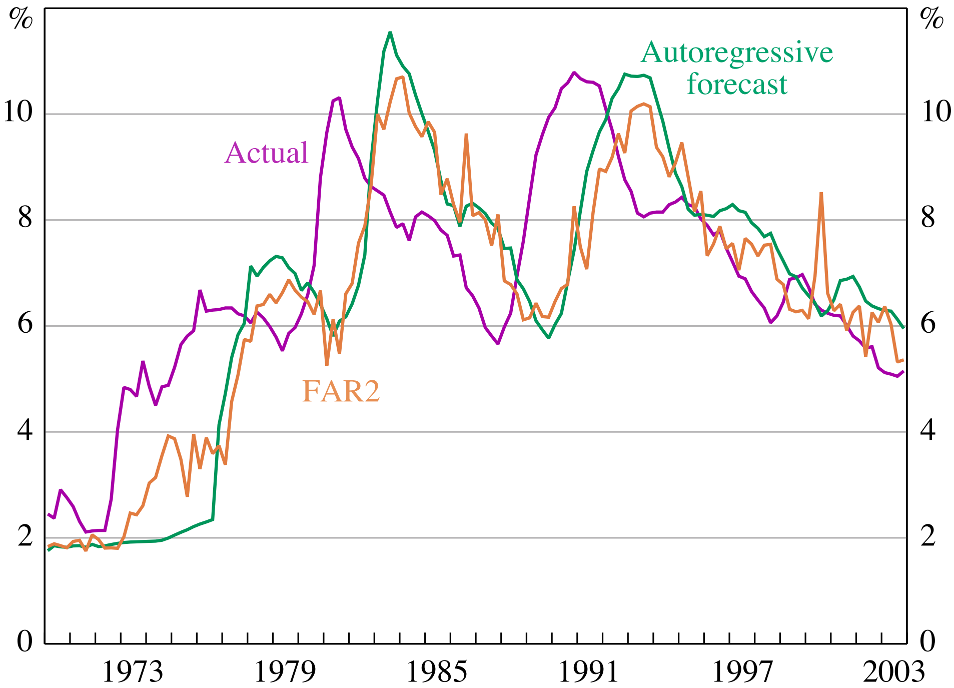
Figure 2: Non-farm GDP Growth Forecasts

Eight-quarter forecasts



5 Interest rates are not included in the data panel as market-determined interest rates are not available for the full sample. Potentially, over a shorter sample, factors from a panel that includes interest rates would be more successful in forecasting building approvals.

Figure 3: Unemployment Rate Forecasts
Eight-quarter forecasts



4.2 Panel Timeliness and Forecast Accuracy

The results in Table 2 demonstrate that factor forecasts can outperform benchmark autoregressive forecasts for many key macroeconomic series. But the breadth of the panel used to generate these factors comes at the expense of including less timely series. As Figure 1 shows, many series have a publication lag of two months or more. In this section we examine how forecast accuracy changes if more timely data are used to generate the factors. To do this, we reproduce the forecasting exercise in Section 4.1 using progressively broader, but less timely, data panels. As outlined in Section 2.3, we include one lag of every series, along with the series that are available at each publication date. This method incorporates the broadest range of up-to-date data. We start with out-of-sample forecasts based on data available 24 days before the end of each quarter. Hence, the panel consists of 57 series: the one-period lag of all 53 series and up-to-date values for the four survey series. We repeat the exercise based on data available at the end of the quarter, allowing us to incorporate an extra three financial market series. We continue this process by moving along the timeline of release dates shown in Figure 1, progressively expanding the number of series included in the panel until it contains 106 series: the one-period lag and up-to-date values for each of the

53 series. With this sequence of out-of-sample forecasts we can then examine how the MSE ratio changes as the forecasts become less timely but the panel uses a more comprehensive set of information.

One important consideration in this exercise is when the base quarter data for the series being forecast become available. For example, the information set used to forecast CPI inflation 20 days after the end of the base quarter will not contain the CPI release for the base quarter, while that used 30 days after the end of the base quarter will. Clearly this adds to the breadth of the panel used to estimate the factors, but it also enables a more up-to-date lag to be included in the autoregressive terms. Our factor forecasts account for this, so that the FAR2 model only contains autoregressive lags 1–2 before the release date of the series being forecast, but contains lags 0–2 after the release date. However, to simplify the interpretation of the change in forecast accuracy as the breadth of the panel changes, we allow the benchmark autoregressive forecast to always use the base quarter release of the series. This means that the denominator of the MSE ratio does not change along with the timeliness of the forecast. Because of this, the MSE ratio before the release of the forecast series does not represent a fair test of forecast accuracy as the factor forecast does not use the base quarter's value of the forecast series while the benchmark autoregressive forecast does.

The MSE ratios for four- and eight-quarter-ahead forecasts are plotted against forecast timeliness – the number of days from the end of the base quarter that the forecast is made – in Figures 4–11. For each figure, moving from left to right presents the MSE ratio when forecasts become less timely, but consequently use more series to estimate the factors.[6] Beyond the release date of the series being forecast (shown as a vertical dashed line), the forecasts also use one extra autoregressive lag as required. For clarity, the FAR2 specification results are not shown in Figures 4–9 since they are very similar to the F2 results. For the unemployment rate and CPI forecasts, the FAR2 specification substantially

6 The ACCI-Westpac survey data are released before the previous quarter's national accounts. In forecasting the national account variables we include the previous quarter's national accounts data in the information used. In effect this means the first forecast would be made around one week after the ACCI-Westpac data are actually released. Similarly, the building approvals release has a publication lag of around 108 days, meaning that at the end of the base quarter it is not yet available for the previous quarter. Despite this we include it in our panel for completeness. Excluding it does not significantly alter the results.

outperforms the F2 specification (as discussed below and in Section 4). So for these two series, the FAR2 results are shown in place of the F2 results (Figures 10–11).

The forecasting performance of the factor models is similar for three of the national accounts series – GDP, non-farm GDP and private final demand – and building approvals (Figures 4–7). In each case the MSE ratios for both the F2 model and FAR-BIC model are always less than 1, demonstrating that the factor forecast is more accurate than the autoregressive forecast. Note that this is even more impressive in light of the fact that the autoregressive forecast uses the most recent quarter's value for the forecast series prior to its release date, while the factor forecast does not. For these series, the MSE ratio tends to be around 0.8 (though with a range from around 0.65 to 0.85), indicating that the factor forecasts are around 20 per cent more accurate than the autoregressive forecast.

Figure 4: Forecast Accuracy by Timeliness of Forecasts

GDP growth

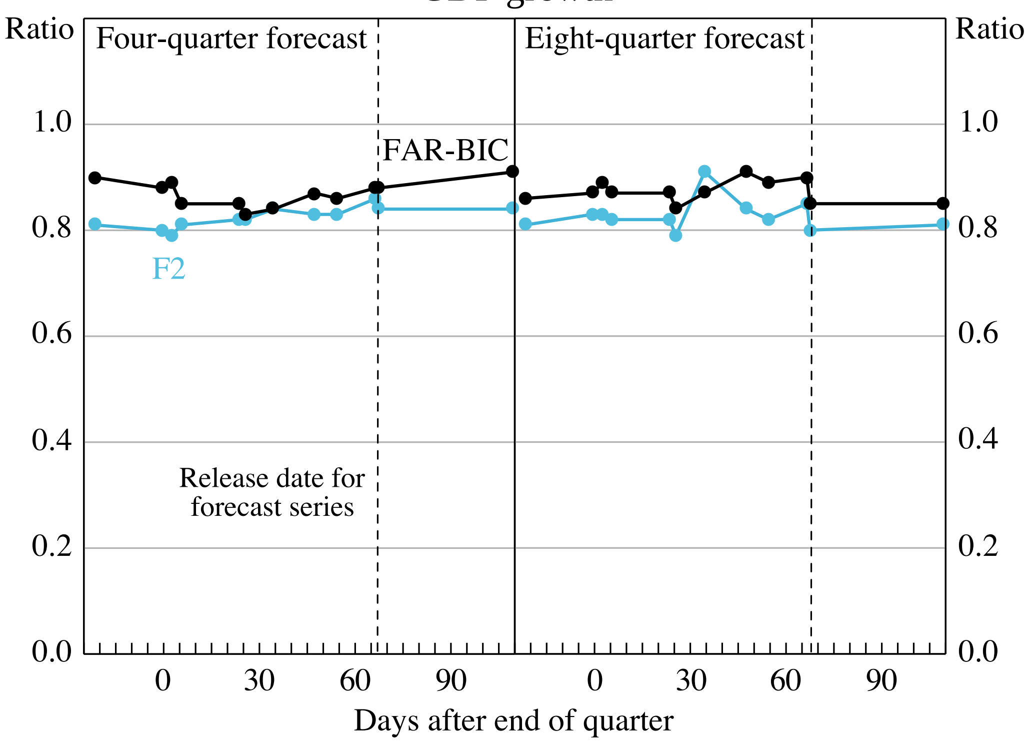


Figure 5: Forecast Accuracy by Timeliness of Forecasts

Non-farm GDP growth

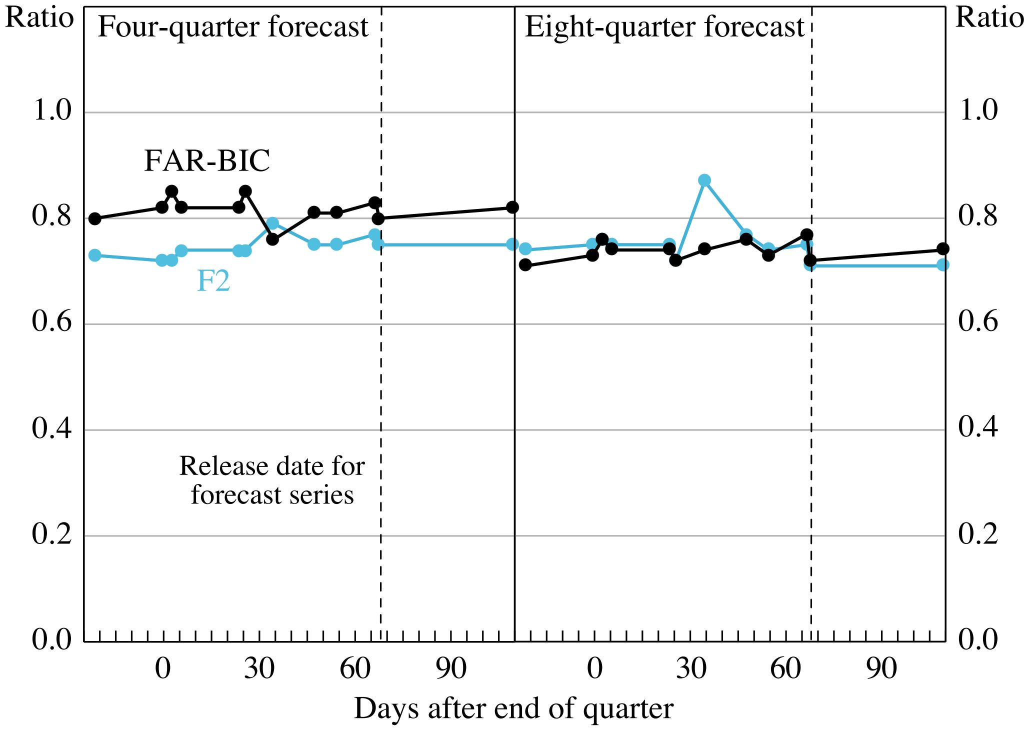


Figure 6: Forecast Accuracy by Timeliness of Forecasts

Private final demand growth

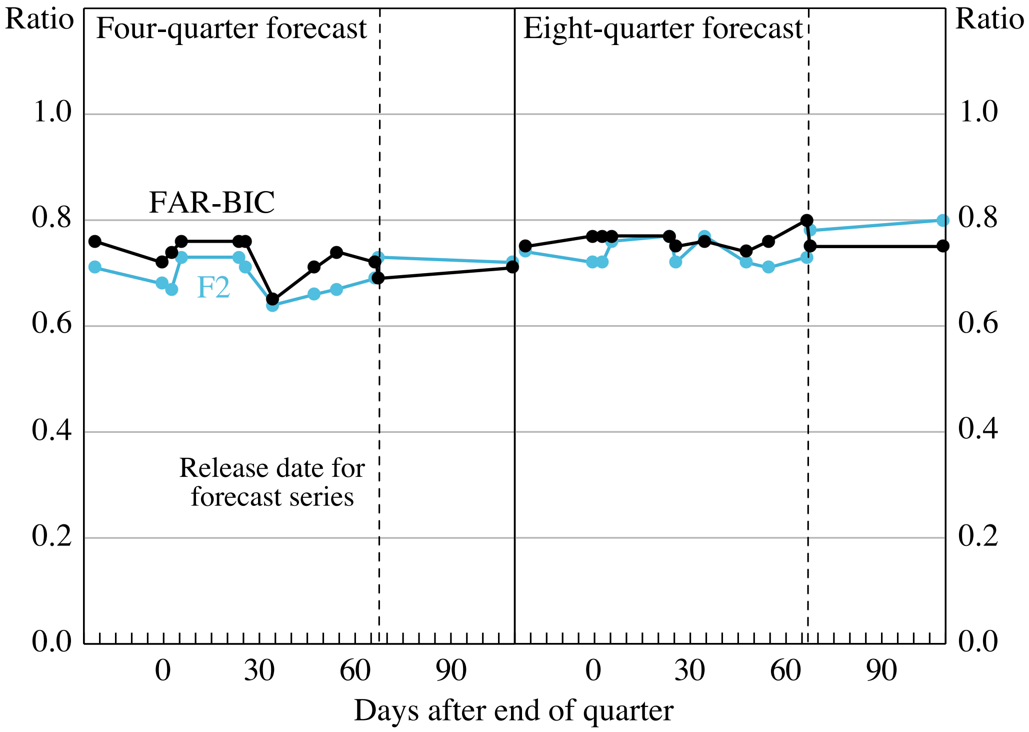


Figure 7: Forecast Accuracy by Timeliness of Forecasts

Building approvals growth

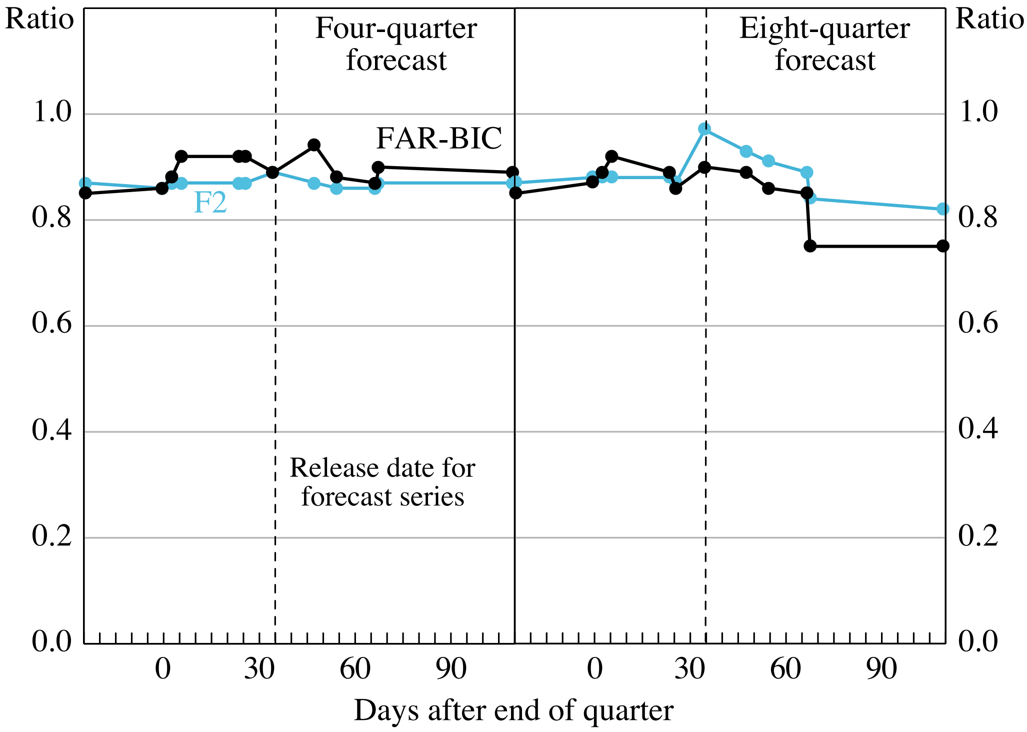


Figure 8: Forecast Accuracy by Timeliness of Forecasts

Household final consumption expenditure growth

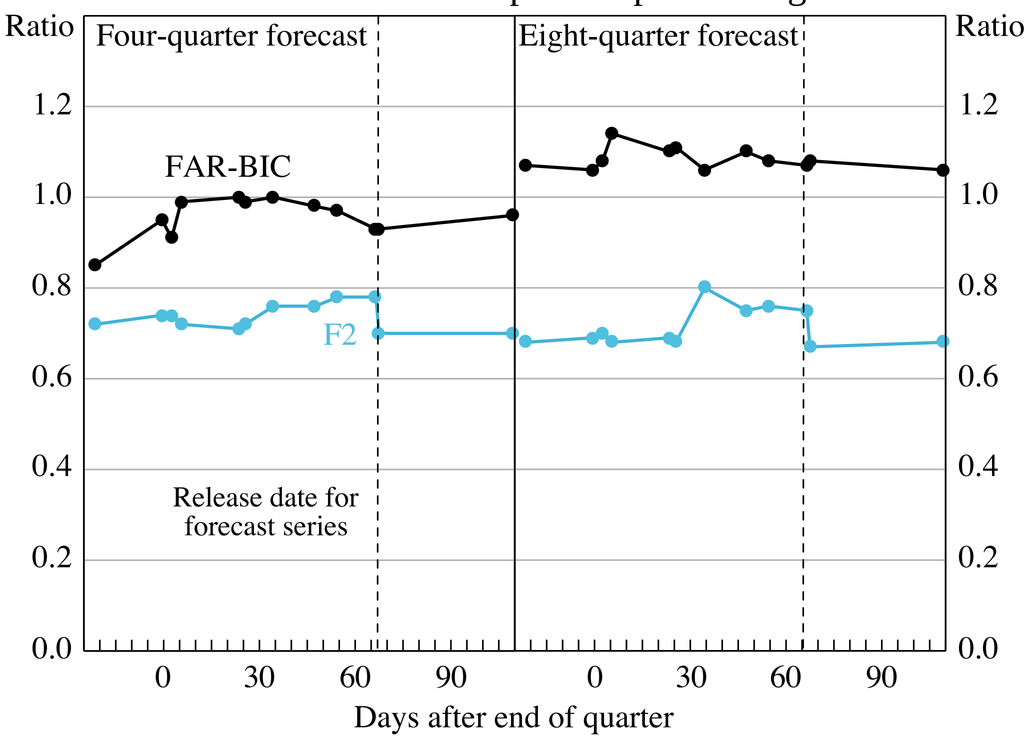


Figure 9: Forecast Accuracy by Timeliness of Forecasts

Employment growth

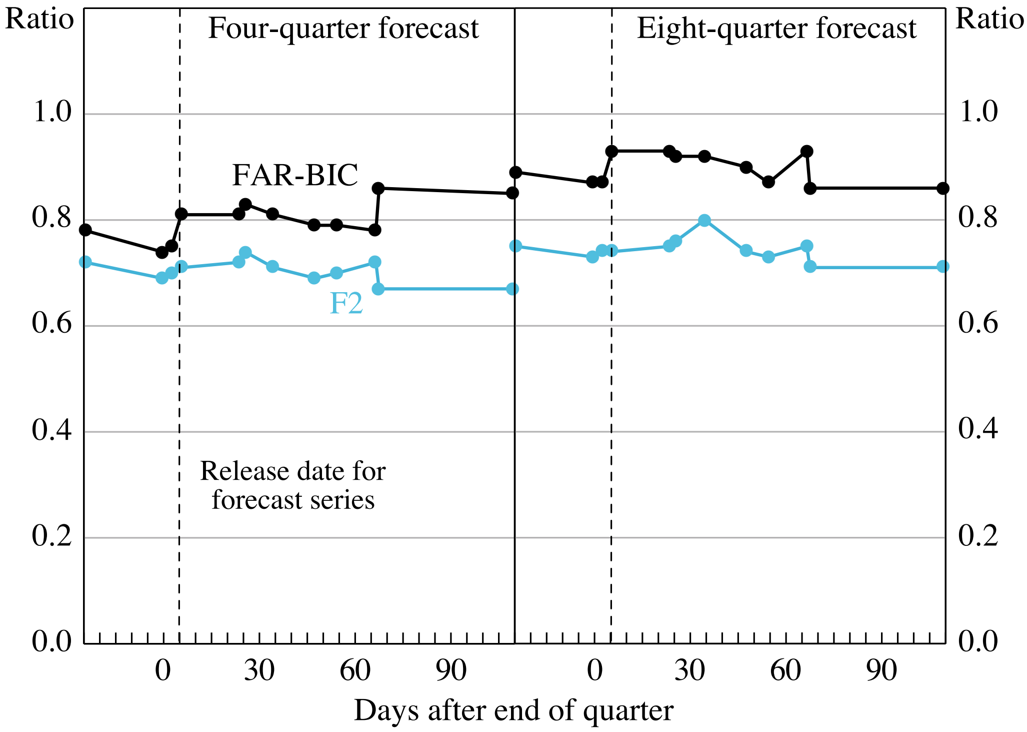


Figure 10: Forecast Accuracy by Timeliness of Forecasts

Unemployment rate

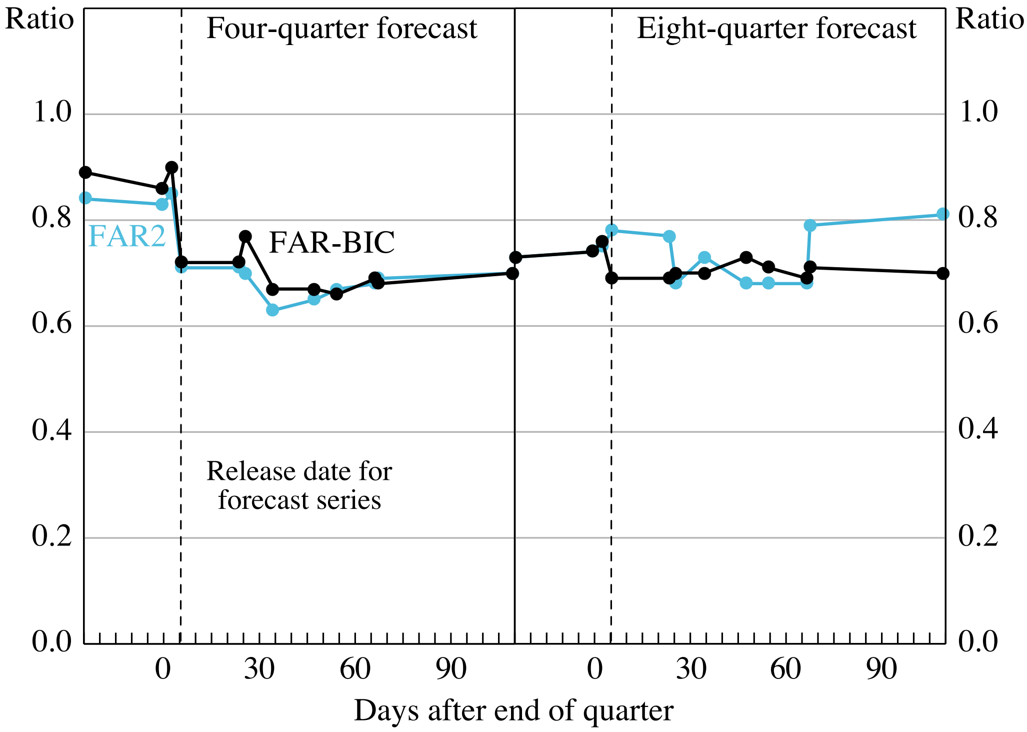
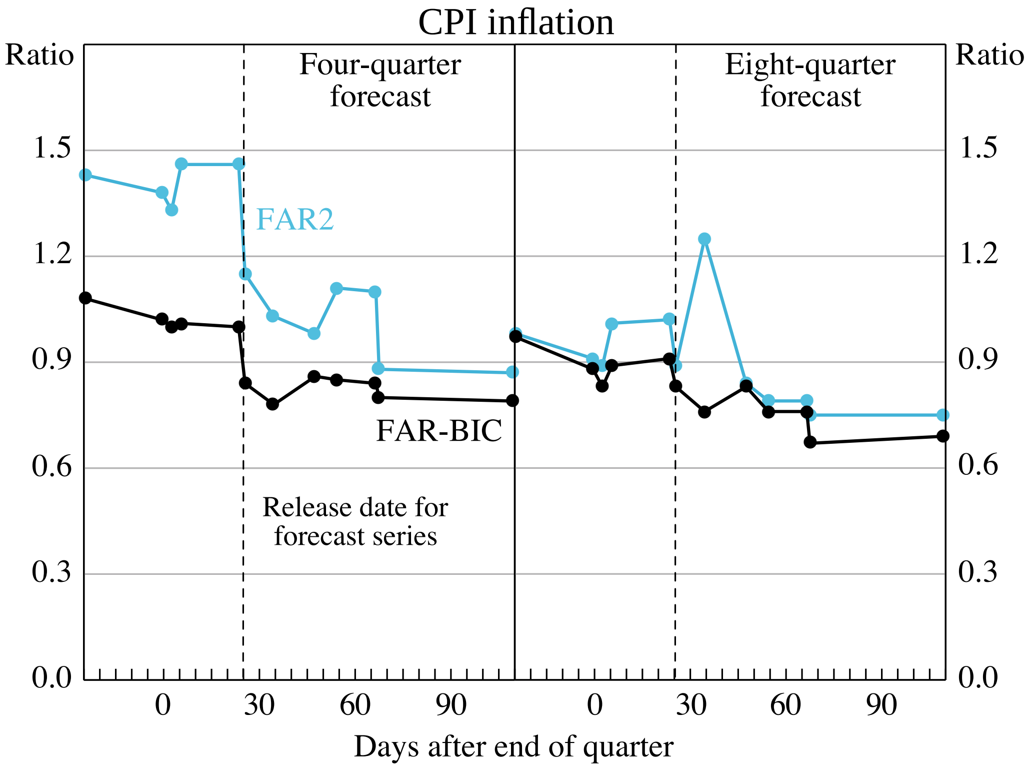


Figure 11: Forecast Accuracy by Timeliness of Forecasts



The other striking feature for all four of these series, at both the four- and eight-quarter horizons, is that the MSE ratios do not have a discernable trend. This indicates that the accuracy of the forecasts does not change substantially (for better or worse) as the data panel is expanded to incorporate the additional series that become available. For these series there is no deterioration in forecast accuracy when forecasts become more timely. This is perhaps not surprising as Gillitzer *et al* (2005) showed that the factors using a similar data panel to that employed here are highly persistent. Since the factors derived from a given quarter's data are very similar to the factors derived from the following quarter, the forecasts based on those factors, and so their errors, are also very similar.

For the other four series – household final consumption expenditure, employment, the unemployment rate and CPI inflation – there are greater differences in the forecast accuracy, across factor models as well as with timeliness (Figures 8 and 11).

For household final consumption expenditure and employment, the simpler two-factor models (F2 and FAR2) are consistently more accurate than the more complex FAR-BIC model. Recall that at each iteration the FAR-BIC model chooses the number of factors and lags of the forecast variable to include in the forecast equation. It is not the inclusion of autoregressive terms that leads to this deterioration in forecast performance; in all three cases the FAR-BIC underperforms relative to the FAR2, which includes autoregressive terms (not shown). Rather, the deterioration in forecast performance is due to the model changing the number of factors at each forecast iteration. The only series for which the FAR-BIC model outperforms the simpler models is CPI inflation. This suggests that for most series, the FAR-BIC model has a tendency to over-fit the data in the period used to estimate the forecasting equations. In contrast, there appear to have been greater structural changes to the inflation process over the long sample, meaning that the changing structure of the FAR-BIC model produces more accurate forecasts. Overall, these results demonstrate that there are benefits to a parsimonious factor model that keeps the number of factors constant.

For both employment growth and the unemployment rate, the MSE ratios for both factor models are always less than one (as they are for GDP, non-farm GDP, private final demand and building approvals), demonstrating that these models produce more accurate forecasts than the autoregressive models. This also applies for most of the forecasts of household final consumption expenditure, with the exception of the FAR-BIC forecasts at an eight-quarter horizon. For CPI inflation at the four-quarter horizon, the MSE ratio for both factor models is initially greater than one, indicating that the factor model forecasts are less accurate than the simple autoregressive model. However, as the forecasts become less timely, and so the data panel used to calculate the forecasts includes more information, the forecast accuracy of the factor models improves and eventually exceeds that of the autoregressive forecasts. The factor model forecasts are generally more accurate relative to the autoregressive forecasts at the eight-quarter horizon than at the four-quarter horizon. These longer horizon forecasts for CPI inflation also tend to become more accurate as the data panel expands.

For two of the series, the unemployment rate and CPI inflation, there is a sharp improvement in the factor forecasts when the base quarter's value of each of these series is included in the information set used for their forecasts; that is, the MSE ratio steps down at the vertical dashed line. As discussed in Section 4, these two series have had long cycles and appear to have experienced considerable structural change. As a result, the FAR2 model, which includes autoregressive lags along with the two factors, substantially outperforms the simpler F2 model that excludes autoregressive lags. The usefulness of autoregressive lags apparently carries through to those lags being more timely, hence the step down in the MSE ratio for both factor models when the base quarter's lag becomes available.

There is one caveat to our observation: that, for most series, forecast accuracy does not improve markedly with broader but less timely panels. For four of the series – household final consumption expenditure, employment, CPI inflation and building approvals – there is some evidence of a small improvement in forecast accuracy of the factor models, as indicated by the step down in the MSE ratio when the national accounts series are included in the panel 68 days after the end of the base quarter.

5. Conclusion

This paper shows that factor-based forecasts can outperform standard time-series benchmarks for key Australian macroeconomic series, as has been found for many other countries. This is perhaps not surprising. Using a relatively large number of series can produce a less noisy estimate of the current state of the economy and also makes the forecasts less susceptible to structural change in any given explanatory variable.

We find almost uniformly that simple models which use a fixed number of factors outperform more complex models that select a different number of factors at each forecast iteration. Further, for most series, the simplest model we present (which excludes lags of the series being forecast) tends to outperform those that include lags. There are two important exceptions to this: forecasts of CPI inflation and the unemployment rate need to include autoregressive lags to account for structural change and the long cycles in these series over our sample period of 45 years.

The use of a broad data panel to estimate the factors may enhance forecast accuracy but at the cost of including series with late publication dates, so resulting in less timely forecasts. We conduct an out-of-sample forecasting exercise that iteratively uses less timely data panels, which contain more information to estimate the factors, in order to assess the extent of this possible trade-off. With the exception of CPI inflation, the forecasts do not become more accurate when they utilise a broader but less timely selection of series. While this is an important result, it is probably not surprising as the factors, which capture economic cycles, are highly persistent. Consequently, the factors derived from adjacent quarters' data tend to be very similar and so are the forecasts. So while factor forecasts have large data requirements, we show that these should not prevent their practical use in a policy setting in which timely forecasts are needed.

Appendix A: Data Panel

Table A1: Data Panel *(continued next page)*

	Code	Time
National accounts		
GDP, sa	5	68
Non-farm GDP, sa	5	68
Public final demand: ex-transfers, sa	5	68
Private final demand: ex-transfers, sa	5	68
Private gross fixed capital formation (GFCF): total, sa	5	68
Gross operating surplus:		
Financial corporations, sa	5	68
Private non-financial corporations, sa	5	68
Public non-financial corporations, sa	5	68
Household final consumption expenditure:		
Total, sa	5	68
Cigarettes & tobacco, sa	5	68
Alcoholic beverages, sa	5	68
Clothing & footware, sa	5	68
Food, sa	5	68
Furnishing & household equipment, sa	5	68
Purchase of vehicles, sa	5	68
Rent & other dwelling services, sa	5	68
Hotels, cafes & restaurants, sa	5	68
Transport services, sa	5	68
Private GFCF: dwellings:		
Alterations and additions, sa	5	68
New and used, sa	5	68
Private non-farm inventories to sales ratio, sa	1	68
Labour market *(continued next page)*		
Employment:		
Males, sa	5	6
Females, sa	5	6
Total, sa	5	6
Unemployment rate, sa	1	6
Labour productivity: heads, sa	5	68
Real unit labour costs, sa	5	68

Table A1: Data Panel *(continued)*

	Code	Time
Labour market *(continued)*		
Average weekly earnings:		
Males, sa	5	48
Females, sa	5	48
ACCI-Westpac surveys:		
Capacity utilisation, net balance, nsa	1	–24
General business situation, next 6mths net balance, nsa	1	–24
Output actual, change in past 3mths net balance, nsa	1	–24
Output expected, change in next 3mths net balance, nsa	1	–24
Building		
Commencements: total new houses and flats excl conversions, number, sa	5	110
Completed: total new houses and flats excl conversions, number, sa	5	110
Building approvals:		
Private new houses and flats, number, sa	5	35
Public new houses and flats, number, sa	5	35
Total new houses and flats, number, sa	5	35
Internal trade		
Motor vehicle registrations, sa	5	24
Overseas transactions		
Current account: excluding RBA gold and frigate, as a share of GDP, sa	1	68
Services imports, sa	5	67
Services exports, sa	5	67
Goods imports, sa	5	67
Goods exports, sa	5	67
Prices		
Consumer price index, nsa	5	26
ABS house prices, nsa	5	55
GDP deflator, sa	5	68
Export price index: goods and services credits IPD, sa	5	67
Import price index: goods and services debits IPD, sa	5	67
Financial data		
Oil prices	5	0
Commodity price index	5	3
Share prices	5	0
Real trade-weighted exchange rate	5	0

Notes: Code represents the transformation made to the data series: 1 indicates no transformation, 5 indicates a log-difference, as in Stock and Watson (2002b). Time is the number of days after the end of the quarter that the data series is published, based on release dates for 2006:Q1.

Appendix B: Alternative Factor Estimates

An alternative technique for estimating the factors has been developed by Forni *et al* (2005) (FHLR). Their methodology takes into account the possibility of leading and lagging relationships of the series in the data panel, and so is referred to as being 'dynamic' (while the technique we use in the paper is referred to as 'static'). The estimation of the dynamic model is more complex than that of the static model, as the factors are estimated in the frequency domain rather than the time domain. While the steps involved in the two estimation procedures differ, conceptually they are still closely related. As Stock and Watson (2006) note, while principal components of the static approach has a least squares interpretation, the dynamic approach has a weighted least squares interpretation. Boivin and Ng (2005) succinctly describe the steps involved in estimating the dynamic factors and provide a comparison of the static and dynamic methods. We use their notation in this section.

We present two different ways of forecasting with the dynamic factors. The first, denoted FHLR-DU, estimates the dynamic factors and then uses them in the forecasting equation, Equation (2), as done with the static factors in Section 2.2. The second, non-parametric, approach produces a forecast directly by projecting forward the common component for each forecast series, and is denoted FHLR-DN. Which technique produces more accurate forecasts is ultimately an empirical issue. Table B1 reproduces the results from Table 2 in Section 4 along with equivalent results for these two dynamic factor methods.

The difference in forecast accuracy between the Stock and Watson static (SW) and FHLR-DU forecasting techniques is typically less than 5 per cent relative to the autoregressive benchmark for all forecasting model specifications. There is some evidence that the SW model performs better at the two- and four-quarter horizons, and the FHLR-DU model at the eight-quarter horizon for the series we forecast, but the difference in forecast accuracy is very small.

Aside from the unemployment rate and CPI series, there is also little difference in forecast accuracy between the SW and FHLR-DN techniques. For the unemployment rate and CPI series, the FHLR-DN techniques forecasts are substantially worse than those made using the SW technique, most likely because the FHLR-DN technique cannot capture the persistence in these series, which the SW and FHLR-DU forecasting models can with the incorporation of autoregressive terms in the forecasting equation. Indeed, the forecast accuracy of the SW technique is comparable to, if not worse than, the FHLR-DN technique for the unemployment rate and CPI series when autoregressive terms are not included in the forecasting equation.

These results confirm the findings of other authors that the simpler SW technique can be used to generate forecasts that are typically at least as accurate as those made using either of the more complicated FHLR techniques. Other forecasts, not reported here, show that the results in the text indicating forecast accuracy conditional on the timeliness of the data are qualitatively similar using the FHLR-DU technique. Because the FHLR-DN technique requires the forecast series to be in the data panel at all times, it is not suitable for our evaluation of forecast accuracy conditional on the timeliness of the data panel.

Table B1: Forecasting Performance *(continued next page)*

Ratio of mean-squared forecast error of candidate model to an autoregressive forecast

Model		**Horizon**					
		Two quarters		Four quarters		Eight quarters	
GDP growth							
SW	*F2*	0.92	(0.12)	0.85	(0.14)	0.83*	(0.13)
	FAR2	0.96	(0.17)	0.89	(0.17)	0.84	(0.14)
	FAR-BIC	1.10	(0.14)	1.02	(0.16)	0.78*	(0.14)
FHLR-DU	*F2*	0.99	(0.11)	0.91	(0.13)	0.79*	(0.14)
	FAR2	1.00	(0.16)	0.94	(0.16)	0.79*	(0.15)
	FAR-BIC	1.11	(0.17)	1.06	(0.19)	0.75*	(0.16)
FHLR-DN		0.88**	(0.07)	0.86*	(0.09)	0.74**	(0.14)
Non-farm GDP growth							
SW	*F2*	0.78**	(0.12)	0.76*	(0.15)	0.73*	(0.17)
	FAR2	0.86	(0.15)	0.87	(0.14)	0.74*	(0.17)
	FAR-BIC	0.95	(0.14)	0.93	(0.14)	0.78*	(0.15)
FHLR-DU	*F2*	0.86*	(0.10)	0.81*	(0.13)	0.68**	(0.18)
	FAR2	0.91	(0.14)	0.92	(0.12)	0.68**	(0.18)
	FAR-BIC	0.84	(0.13)	0.83	(0.15)	0.70**	(0.18)
FHLR-DN		0.80**	(0.09)	0.80**	(0.11)	0.71**	(0.16)

Table B1: Forecasting Performance *(continued next page)*

Ratio of mean-squared forecast error of candidate model to an autoregressive forecast

Model		Horizon					
		Two quarters		Four quarters		Eight quarters	
Private final demand growth							
SW	*F2*	0.82*	(0.12)	0.71**	(0.15)	0.77*	(0.15)
	FAR2	0.78**	(0.13)	0.67**	(0.17)	0.73*	(0.17)
	FAR-BIC	0.81*	(0.13)	0.74*	(0.18)	0.80	(0.16)
FHLR-DU	*F2*	0.86	(0.11)	0.77**	(0.12)	0.73**	(0.14)
	FAR2	0.80**	(0.12)	0.71**	(0.15)	0.74**	(0.15)
	FAR-BIC	0.83*	(0.12)	0.63**	(0.19)	0.72*	(0.18)
FHLR-DN		0.79**	(0.09)	0.81**	(0.10)	0.78**	(0.13)
Household final consumption expenditure growth							
SW	*F2*	0.90*	(0.07)	0.69***	(0.11)	0.64**	(0.16)
	FAR2	0.97	(0.08)	0.69***	(0.11)	0.65**	(0.16)
	FAR-BIC	1.20	(0.26)	0.97	(0.21)	1.00	(0.26)
FHLR-DU	*F2*	0.93	(0.07)	0.72***	(0.11)	0.64***	(0.15)
	FAR2	0.97	(0.07)	0.73***	(0.11)	0.66**	(0.15)
	FAR-BIC	1.20	(0.29)	0.97	(0.22)	0.89	(0.16)
FHLR-DN		0.89**	(0.05)	0.79***	(0.08)	0.67**	(0.14)
Employment growth							
SW	*F2*	0.76**	(0.14)	0.71**	(0.16)	0.77*	(0.17)
	FAR2	0.78**	(0.13)	0.76*	(0.17)	0.80	(0.18)
	FAR-BIC	0.82*	(0.12)	0.84	(0.14)	0.81	(0.17)
FHLR-DU	*F2*	0.78**	(0.13)	0.73**	(0.15)	0.75*	(0.17)
	FAR2	0.78**	(0.13)	0.77*	(0.15)	0.81	(0.18)
	FAR-BIC	0.85	(0.12)	0.81	(0.15)	0.77	(0.18)
FHLR-DN		0.83**	(0.10)	0.80*	(0.11)	0.79**	(0.12)
Unemployment rate							
SW	*F2*	16.02	(4.06)	4.08	(2.65)	1.80	(0.61)
	FAR2	0.67*	(0.16)	0.71*	(0.16)	0.76*	(0.13)
	FAR-BIC	0.58*	(0.19)	0.70*	(0.17)	0.71*	(0.15)
FHLR-DU	*F2*	16.33	(37.92)	4.06	(2.88)	1.68	(0.55)
	FAR2	0.71*	(0.15)	0.73*	(0.15)	0.70*	(0.14)
	FAR-BIC	0.59*	(0.19)	0.70*	(0.18)	0.68*	(0.16)
FHLR-DN		6.14	(5.93)	2.22	(0.78)	1.39	(0.24)

Table B1: Forecasting Performance *(continued)*

Ratio of mean-squared forecast error of candidate model to an autoregressive forecast

Model		**Horizon**					
		Two quarters		Four quarters		Eight quarters	
Consumer price inflation							
SW	*F2*	1.78	(0.56)	1.63	(0.40)	1.14	(0.22)
	FAR2	0.89	(0.12)	0.84	(0.16)	0.71*	(0.20)
	FAR-BIC	1.01	(0.20)	0.78	(0.21)	0.63*	(0.28)
FHLR-DU	*F2*	1.92	(0.69)	1.61	(0.36)	1.14	(0.19)
	FAR2	1.08	(0.16)	0.88	(0.19)	0.74	(0.23)
	FAR-BIC	1.02	(0.19)	0.86	(0.22)	0.78	(0.27)
FHLR-DN		2.22	(0.80)	1.57	(0.41)	1.20	(0.17)
Building approvals growth							
SW	*F2*	1.09	(0.10)	1.04	(0.09)	0.94	(0.11)
	FAR2	1.01	(0.08)	1.04	(0.09)	0.98	(0.12)
	FAR-BIC	0.98	(0.09)	0.98	(0.09)	1.01	(0.11)
FHLR-DU	*F2*	1.19	(0.11)	1.11	(0.10)	0.93	(0.10)
	FAR2	1.07	(0.08)	1.05	(0.09)	1.00	(0.11)
	FAR-BIC	0.89	(0.09)	0.90	(0.09)	0.98	(0.10)
FHLR-DN		1.05	(0.06)	1.03	(0.06)	1.10	(0.09)

Notes: SW indicates factors estimated using the technique of Stock and Watson (1999; 2002a; 2002b). FHLR indicates factors estimated using the technique of Forni, Hallin, Lippi and Reichlin (2005). In the notation of Boivin and Ng (2005), DU indicates 'dynamic unrestricted' factors, while DN indicates 'dynamic non-parametric factors'. Model F2 includes two factors and no lags of the forecast variable. Model FAR2 includes two factors and up to four lags of the forecast variable, selected at each iteration using the BIC ($0 \leq p \leq 3$). Model FAR-BIC uses the BIC to select both the number of factors (up to six, $1 \leq k \leq 6$) and the number of lags of the forecast variable (up to three, $0 \leq p \leq 3$) at each iteration. Numbers in parentheses are robust standard errors calculated using the delta method. Ratios significantly less than 1 at the 1, 5 and 10 per cent confidence levels are indicated by ***, ** and *.

References

Angelini E, J Henry and R Mestre (2001), 'Diffusion Index-Based Inflation Forecasts for the Euro Area', in *Empirical Studies of Structural Changes and Inflation*, BIS Papers, 3, pp 109–138.

Artis MJ, A Banerjee and M Marcellino (2005), 'Factor Forecasts for the UK', *Journal of Forecasting*, 24(4), pp 279–298.

Banerjee A, M Marcellino and I Masten (2005), 'Forecasting Macroeconomic Variables for the New Member States of the European Union', European Central Bank Working Paper No 482.

Boivin J and S Ng (2005), 'Understanding and Comparing Factor-Based Forecasts', *International Journal of Central Banking*, 1(3), pp 117–151.

Boivin J and S Ng (2006), 'Are More Data Always Better for Factor Analysis?', *Journal of Econometrics*, 132(1), pp 169–194.

Brave S and JDM Fisher (2004), 'In Search of a Robust Inflation Forecast', Federal Reserve Bank of Chicago *Economic Perspectives*, 28(4), pp 12–31.

Forni M, M Hallin, M Lippi and L Reichlin (2005), 'The Generalized Dynamic Factor Model: One-Sided Estimation and Forecasting', *Journal of the American Statistical Association*, 100(471), pp 830–840.

Gavin WT and KL Kliesen (2006), 'Forecasting Inflation and Output: Comparing Data-Rich Models with Simple Rules', Federal Reserve Bank of St. Louis Working Paper No 2006-054A.

Gillitzer C, J Kearns and A Richards (2005), 'The Australian Business Cycle: A Coincident Indicator Approach', in C Kent and D Norman (eds), *The Changing Nature of the Business Cycle*, Proceedings of a Conference, Reserve Bank of Australia, Sydney, pp 262–307.

Gosselin M-A and G Tkacz (2001), 'Evaluating Factor Models: An Application to Forecasting Inflation in Canada', Bank of Canada Working Paper No 2001-18.

Matheson T (2005), 'Factor Model Forecasts for New Zealand', Reserve Bank of New Zealand Discussion Paper No DP2005/01.

Moser G, F Rumler and J Scharler (2004), ‘Forecasting Austrian Inflation’, Oesterreichische Nationalbank Working Paper No 91.

Schneider M and M Spitzer (2004), ‘Forecasting Austrian GDP Using the Generalized Dynamic Factor Model’, Oesterreichische Nationalbank Working Paper No 89.

Stock JH and MW Watson (1999), ‘Forecasting Inflation’, *Journal of Monetary Economics*, 44(2), pp 293–335.

Stock JH and MW Watson (2002a), ‘Forecasting Using Principal Components from a Large Number of Predictors’, *Journal of the American Statistical Association*, 97(460), pp 1167–1179.

Stock JH and MW Watson (2002b), ‘Macroeconomic Forecasting Using Diffusion Indexes’, *Journal of Business and Economic Statistics*, 20(2), pp 147–162.

Stock JH and MW Watson (2006), ‘Forecasting with Many Predictors’, in G Elliott, CWJ Granger and A Timmermann (eds), *Handbook of Economic Forecasting*, *Volume 1*, Elsevier Science Ltd, North-Holland, pp 515–554.

Watson MW (2003), ‘Macroeconomic Forecasting Using Many Predictors’, in M Dewatripont, LP Hansen and SJ Turnovsky (eds), *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress*, *Volume III*, Cambridge University Press, pp 87–114.